Procedia Computer Science

Volume 80, 2016, Pages 1255–1266

ICCS 2016. The International Conference on Computational Science

Variance-reduced HMM for Stochastic Slow-Fast Systems

Ward Melis[1] and Giovanni Samaey[1]

Department of Computer Science, K.U. Leuven, Celestijnenlaan 200A, 3001 Leuven, Belgium
ward.melis@cs.kuleuven.be

Abstract

We propose a novel variance reduction strategy based on control variables for simulating the averaged equation of a stochastic slow-fast system. In this system, we assume that the fast equation is ergodic, implying the existence of an invariant measure, for every fixed value of the slow variable. The right hand side of the averaged equation contains an integral with respect to this unknown invariant measure, which is approximated by the heterogeneous multiscale method (HMM). The HMM method corresponds to a Markov chain Monte Carlo method in which samples are generated by simulating the fast equation for a fixed value of the slow variable. As a consequence, the variance of the HMM estimator decays slowly. Here, we introduce a variance-reduced HMM estimator based on control variables: from the current time HMM estimation, we subtract a second HMM estimator at the previous time step using the same seed as the current time HMM estimator. To avoid introducing a bias, we add the previously calculated variance-reduced estimator. We analyze convergence of the proposed estimator and apply it to a linear and nonlinear model.

Keywords: multiscale, stochastic, slow-fast, HMM, variance reduction, control variables

1 Introduction

Stochastic differential equations are ubiquitous in a multitude of real-life applications appearing in different scientific domains, such as climate and environmental sciences [1], molecular dynamics [2, 4] and bacterial chemotaxis [11]. Many of these applications contain processes that inherently evolve over multiple time scales. One prototypical example system that was proposed in [4] to analyze convergence behavior of novel multiscale methods is a singularly perturbed slow-fast system in which the slow variable is described deterministically, while the model for the fast variable contains stochastic effects. The general form is as follows:

$$\begin{cases} dx(t) = f(x,y)dt, & x(0) = x_0 \in \mathbb{R} \\ dy(t) = \frac{1}{\varepsilon} g(x,y)dt + \frac{1}{\sqrt{\varepsilon}} \beta(x,y)dW(t), & y(0) = y_0 \in \mathbb{R}, \end{cases} \tag{1}$$

where the scalar quantities $x(t) : [0,T] \to \mathbb{R}$ and $y(t) : [0,T] \to \mathbb{R}$ represent the slow and fast evolving stochastic processes, respectively. The functions $f(x,y), g(x,y) \in \mathbb{R}$ are called the

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2016

doi:10.1016/j.procs.2016.05.497

drift functions and $\beta(x, y) \in \mathbb{R}$ is termed the diffusion function. This nomenclature originates from the effect of these terms in the corresponding Fokker-Planck equation. Furthermore, $W(t) \in \mathbb{R}$ denotes a standard Brownian motion. The parameter $\varepsilon \ll 1$ is a positive small-scale parameter measuring the time scale separation between the fast and slow variable in system (1). In addition, we assume that the fast dynamics is ergodic for every fixed state $\bar{x}$ of the slow variable, implying the existence and uniqueness of an invariant measure [10].

Often, one is only interested in the evolution of the slow variable of system (1) and not in a detailed description of the fast variable. However, the fast dynamics cannot be omitted, since the slow process explicitly depends on the fast variable. Due to the stiffness in system (1), explicit simulation techniques such as the Euler-Maruyama or higher-order Milstein schemes are computationally prohibitive. In the deterministic setting, implicit solvers such as the backward Euler and trapezoidal methods are popular choices for stiff systems allowing much larger time steps than explicit methods. However, when dealing with systems with an invariant measure different from a Dirac distribution centered at the equilibrium state of the fast equation, it is known that implicit methods fail to capture the correct invariant measure, thus introducing a bias [7]. To avoid this, it is required to limit the time step of the implicit method by the time scale separation, thus losing its computational advantage in the stochastic setting.

The difficulties related to direct simulation routines can be remedied by exploiting the time scale separation of system (1): for $\varepsilon \to 0$, the averaging principle [10] yields the following reduced description for the slow variable of system (1):

$$\frac{dX}{dt} = F(X), \qquad F(X) = \int_{\mathcal{Y}} f(X, y) d\mu_X^\infty(y). \tag{2}$$

in which $\mu_X^\infty(y)$ denotes the invariant measure induced by the fast dynamics of system (1) keeping $x = X$ fixed. Equation (2) is known as the *averaged*, *macroscopic* or *reduced* equation for the slow variable. The averaging principle transforms the original stochastic slow-fast system into the deterministic problem (2) written in terms of the slow variable only. Given that the latter is completely deterministic, equation (2) can be simulated using explicit integration routines such as the forward Euler or higher-order Runge-Kutta methods. However, before these schemes can be used, a procedure to calculate $F(X)$ in equation (2) is needed. In general, the invariant measure is not known explicitly. Then, one needs to resort to numerical methods that approximate the integral in (2) without explicit knowledge of the invariant measure, such as the heterogeneous multiscale method (HMM) [12] and the projective integration (PI) method [3].

In essence, the HMM method is a Markov chain Monte Carlo method. The variance of the HMM estimator can be large since samples generated by the Markov chain are correlated. Moreover, the variance decays as σ_f^2/M, with σ_f^2 the variance of f, resulting in very slow convergence. This can be improved by either increasing the sample size M, or by lowering the variance σ_f^2. Since the former choice will lead to a much higher computational cost, we examine the latter in this work. The variance reduction method we propose is based on a control variables technique in which a control process is introduced that is strongly correlated with the original HMM estimator. Consequently, by subtracting these two, corresponding variations greatly diminish leading to a variance-reduced HMM estimator.

The remainder of this paper is structured as follows. In section 2, we introduce the stochastic slow-fast systems that we intend to solve numerically. The HMM framework to efficiently integrate these slow-fast systems is described in section 3. In that section, we also present the variance-reduced HMM framework. Next, in section 4 we study the numerical properties of the proposed variance reduction method. Numerical results are reported in section 5. We conclude in section 6 with a brief discussion and ideas for future work.

2 Slow-Fast System

The general form of slow-fast systems we consider is given in equation (1) and describes the coupled evolution of a scalar slow variable $x(t)$ using a deterministic model, and a scalar fast variable $y(t)$ evolving according to a stochastic model. We point out that the results obtained in this work can easily be extended to cases with more fast or slow variables.

In what follows, we will always assume that the fast dynamics of system (1) is ergodic for all fixed values of the slow variable. Ergodicity implies that the statistical properties of the ensemble of the stochastic process at a fixed time instance and those of one realization of the process over an infinite time interval are the same. Consequently, for an ergodic process, averaging a function with respect to the invariant measure $\mu_X^\infty(y)$ yields the same result as averaging this function over one infinitely long time path of the process:

$$F(X) = \int_{\mathcal{Y}} f(X,y) d\mu_X^\infty(y) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(X, y(t)) dt. \tag{3}$$

This interchange between averages serves as a base for numerical methods avoiding explicit knowledge of the invariant measure $\mu_X^\infty(y)$. Here, we consider the following linear and nonlinear model problems.

Linear system. In the linear setting, system (1) takes on the following form:

$$\begin{cases} dx(t) = \big(\lambda x(t) + p y(t)\big) dt \\ dy(t) = \dfrac{1}{\varepsilon}\big(q x(t) - A y(t)\big) dt + \dfrac{1}{\sqrt{\varepsilon}} dW(t), \end{cases} \tag{4}$$

in which the parameters λ, p, q and A are all real scalars. In addition, to ensure that solutions decay exponentially with time, we require that $\lambda < 0$ and $A \in \left(\dfrac{pq}{-\lambda}, 2\right]$.

For this linear system, the invariant measure of the fast equation with respect to every fixed value X of the slow variable can be calculated analytically as:

$$\mu_X^\infty(y) \sim \mathcal{N}(m_\infty, \sigma_\infty^2), \qquad m_\infty = \frac{q}{A} X, \qquad \sigma_\infty^2 = \frac{1}{2A}, \tag{5}$$

where m_∞ and σ_∞^2 denote the mean and variance of the invariant measure, respectively. Since the invariant measure is known, the integral in equation (2) can also be calculated analytically, yielding:

$$\frac{dX}{dt} = \left(\lambda + \frac{pq}{A}\right) X, \qquad X(0) = x_0, \tag{6}$$

which is a linear ODE with exact solution $X(t) = x_0 \exp((\lambda + pq/A)t)$.

Nonlinear system. As a second example, we consider the following nonlinear stochastic multiscale system from [9]:

$$\begin{cases} dx(t) = -\big(y(t) + y(t)^2\big) dt \\ dy(t) = -\dfrac{1}{\varepsilon}\big(y(t) - x(t)\big) dt + \dfrac{1}{\sqrt{\varepsilon}} dW(t). \end{cases} \tag{7}$$

In this case, the dynamics of the slow variable is nonlinear, while the fast variable is described by a linear Ornstein-Uhlenbeck process. For this process, it is known that the invariant measure $\mu_x^\infty(y)$ is a Gaussian with the following invariant mean m_∞ and variance σ_∞^2:

$$m_\infty = x, \qquad \sigma_\infty^2 = \frac{1}{2}. \tag{8}$$

Since the invariant measure is known, the reduced equation for X can be calculated analytically as:

$$\frac{dX}{dt} = -\left(X + X^2 + \frac{1}{2}\right), \qquad X(0) = x_0, \tag{9}$$

which is a nonlinear ODE with solution $X(t) = -0.5 - 0.5\tan\left(0.5t - \arctan(2x_0 + 1)\right)$.

3 Numerical Method

In this section, we construct a variance-reduced numerical scheme to solve the averaged equation (2). Since this equation is deterministic, any stable explicit ODE solver can be used, such as the forward Euler and Runge-Kutta methods. Here, we will employ the forward Euler (FE) method. To that end, we discretize equation (2) on a uniform time mesh with time step Δt, and $t^n = n\Delta t$. The numerical solution on this mesh is denoted by $\hat{X}^n$. The forward Euler scheme is then given by,

$$\hat{X}^{n+1} = \hat{X}^n + \Delta t \hat{F}(X^n), \qquad \hat{X}^0 = x_0. \tag{10}$$

In equation (10), the function F is replaced by an appropriate estimator $\hat{F}$, since, in general, the integral in equation (2) can not be calculated analytically. Therefore, we first introduce the HMM estimator in section 3.1. Then, we present a novel variance-reduced HMM estimator based on control variables in section 3.2.

3.1 Heterogeneous Multiscale Method (HMM)

The heterogeneous multiscale method [12] bypasses explicit knowledge of the invariant measure in equation (2) by exploiting the ergodicity property given in equation (3): F is calculated by averaging over one infinitely long time path of the fast process of system (1) while keeping the value of the slow variable fixed. As a result, the HMM estimator boils down to a Markov chain Monte Carlo estimator: the integral in equation (2) is approximated by a Monte Carlo method, in which the samples are not drawn from the (unknown) invariant measure, but are instead generated from a Markov chain. This chain is obtained by simulating the fast equation using the explicit Euler-Maruyama scheme, which is the stochastic counterpart of the forward Euler scheme. In that regard, we discretize the fast equation on a uniform time mesh with time step δt. For a given fixed value $\hat{X}^n$ of the slow variable, the numerical solution at time $t^{n,m'} = n\Delta t + m'\delta t$ is denoted by $y^{n,m'}$. The Euler-Maruyama scheme is given by [5],

$$y^{n,m'+1} = y^{n,m'} + \frac{\delta t}{\varepsilon} g(\hat{X}^n, y^{n,m'}) + \sqrt{\frac{\delta t}{\varepsilon}}\beta(\hat{X}^n, y^{n,m'})\xi_n^{m'}, \qquad m' = 0, ..., M-1, \tag{11}$$

in which $(\xi_{m'}^n)_{m'=0}^{M-1}$ is a set of mutually independent samples drawn from the standard normal distribution using a random number generator with seed ω_n. The factor $\sqrt{\delta t}$ appears in the last term of equation (11) due to the scaling property of Brownian motion. The initial condition of

the Euler-Maruyama method is chosen as $y^{0,0} = y_0$, and for all other $n > 0$ as $y^{n,0} = y^{n-1,M-1}$. Then, the samples generated by the Markov chain (11) are approximately distributed according to the desired invariant measure. Since it is more natural to label samples from 1 to M, we use the trivial substitution $m = m' + 1$ as sample index. The HMM estimator at time instance t^n is calculated as follows:

$$\hat{F}_{\text{HMM}}(\hat{X}^n; \omega_n) = \frac{1}{M} \sum_{m=1}^{M} f(\hat{X}^n, y^{n,m}), \tag{12}$$

in which ω_n represents the seed that is used in the random number generator. Since we can only generate finite sample sizes M, the HMM estimator in equation (12) is a random variable.

3.2 Variance-reduced HMM

In this work, we propose a variance-reduced estimator based on the well-known control variables technique, in which the variance of the estimator is reduced by introducing a related control process. The expression of the variance-reduced HMM estimator at time instance t^n using control variables becomes:

$$\bar{F}_{\text{HMM}}(\bar{X}^n) = \hat{F}_{\text{HMM}}(\bar{X}^n; \omega_n) - \left(\hat{F}_{\text{HMM}}(\bar{X}^{n-1}; \omega_n) - \bar{F}_{\text{HMM}}(\bar{X}^{n-1}) \right), \tag{13}$$

in which an overbar denotes a variance-reduced estimator. The first term $\hat{F}_{\text{HMM}}(\bar{X}^n; \omega_n)$ in equation (13) coincides with the classical HMM estimator without variance reduction for the slow variable $\bar{X}^n$ using seed ω_n. The second term $\hat{F}_{\text{HMM}}(\bar{X}^{n-1}; \omega_n)$ represents another HMM estimation without variance reduction for the slow variable $\bar{X}^{n-1}$ at the previous time instance. However, this term uses the exact same seed ω_n as in the first term. The last term $\bar{F}_{\text{HMM}}(\bar{X}^{n-1})$ is the variance-reduced HMM estimator previously calculated.

The proposed estimator exploits information on the noise of the estimator at a previous time instance to improve current time instance estimation. The difference between $\hat{F}_{\text{HMM}}(\bar{X}^{n-1}; \omega_n)$ and $\bar{F}_{\text{HMM}}(\bar{X}^{n-1})$ between brackets in equation (13) corresponds to estimator noise. Since the same seed ω_n is used in the first term and noise calculation at the previous time instance, we can assume this noise resembles the HMM estimator noise at the current time instance.

The proposed technique can also be viewed from the following perspective. The first and second estimator terms in equation (13) are strongly correlated, as they employ the same seed ω_n. Consequently, the variations of both estimators will be similar. By subtracting these two terms, we attempt to cancel corresponding variations. To avoid introducing a bias, subtracting the previous variance-reduced estimator from the obtained difference is required. In conclusion, the first and second terms in equation (13) are strongly correlated to reduce the variance, while the terms between brackets have the same expected value, so no bias is introduced in the variance-reduced estimator. The variance reduction method is illustrated in figure 1.

The variance-reduced estimator is used in practice as described below. Given the initial values $\bar{X}^0 = x_0$ and y_0 for the slow and fast variables, respectively, $\bar{X}^1$ is calculated by taking one forward Euler step:

$$\bar{X}^1 = \bar{X}^0 + \Delta t \bar{F}_{\text{HMM}}(\bar{X}^0), \tag{14}$$

where the initial calculation of $\bar{F}_{\text{HMM}}(\bar{X}^0)$ is either the exact solution $F(\bar{X}^0)$ or a more accurately estimated value of (12) achieved through using many more samples. Then, all other values of the slow variable are obtained using the forward Euler scheme

$$\bar{X}^{N+1} = \bar{X}^N + \Delta t \bar{F}_{\text{HMM}}(\bar{X}^N), \tag{15}$$

together with the variance-reduced version of the estimator given in (13).

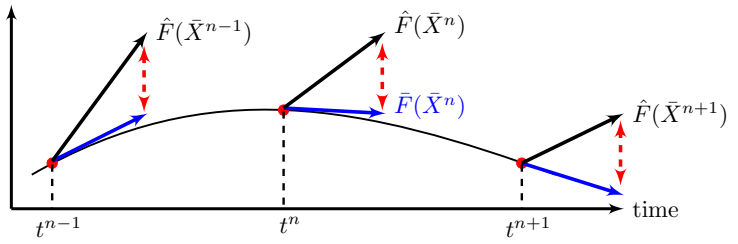

Figure 1: Sketch of the proposed variance reduction technique by control variables. Black arrows denote the estimated value of F by HMM, while blue arrows depict the variance-reduced HMM estimator. The red dashed arrows indicate that the noise – the difference between black and blue arrows – is similar on consecutive time instances (color online).

4 Numerical Properties

In this paper, we only state the main results. A detailed analysis and proofs of the reported results are postponed to a further publication [8]. The assumptions and approximations underpinning the variance-reduced HMM estimator lead to different sources of error. First, when exploiting the ergodicity property (3), we can only simulate the fast dynamics over a finite time interval $[0, \tau]$ with $\tau = M\delta t$, which leads to a first error. Second, we introduce an error by replacing the (finite) time integral by a finite Riemann sum, in which the fast variable is evaluated at discrete time instances. Third, since the exact solution of the fast equation is not known explicitly, a time discretization method is used to approximate the solution of this equation, which leads to a discretization error. The first and third error sources are deterministic and belong to the class of systematic errors. The second error is stochastic, since we are taking a finite sum of a stochastic process, thus residing in the class of statistical errors. A quantity that captures the error of an estimator is the mean squared error (MSE), which satisfies the following equality:

$$\mathbb{E}\left[|\bar{F}_{\mathrm{HMM}}(x) - F(x)|^2\right] = \left(\tilde{F}_{\mathrm{HMM}}(x) - F(x)\right)^2 + \mathbb{E}\left[|\bar{F}_{\mathrm{HMM}}(x) - \tilde{F}_{\mathrm{HMM}}(x)|^2\right], \tag{16}$$

in which we used the following shorthand notation for the expected value of an estimator $\theta(x)$:

$$\tilde{\theta}(x) = \mathbb{E}[\theta(x)]. \tag{17}$$

Here, $\theta(x)$ is either the original or variance-reduced HMM estimator. The first term on the right hand side of equation (16) is the estimator bias squared, while the second term corresponds to the estimator variance. Equation (16) allows for a clear distinction between the systematic and statistical errors of the estimator.

Working out the recursion in equation (13), the variance-reduced HMM estimator at time instance t^N can be written as follows:

$$\begin{aligned}\bar{F}_{\mathrm{HMM}}(\bar{X}^N) = \hat{F}_{\mathrm{HMM}}(\bar{X}^N;\omega_N) + \sum_{n=1}^{N-1}\left(\hat{F}_{\mathrm{HMM}}(\bar{X}^n;\omega_n) - \hat{F}_{\mathrm{HMM}}(\bar{X}^n;\omega_{n+1})\right) \\ - \left(\hat{F}_{\mathrm{HMM}}(\bar{X}^0;\omega_1) - \bar{F}_{\mathrm{HMM}}(\bar{X}^0)\right).\end{aligned} \tag{18}$$

In the following two sections, we look at the convergence of the estimator and pathwise convergence.

4.1 Estimator Convergence

First, we examine the bias of the variance-reduced HMM estimator at time instance t^N. Therefore, we take the expectation of both sides of equation (18) over repeated experiments while keeping M and δt fixed:

$$\mathbb{E}[\bar{F}_{\text{HMM}}(\bar{X}^N)] = \mathbb{E}[\hat{F}_{\text{HMM}}(\bar{X}^N;\cdot)] - \left(\mathbb{E}[\hat{F}_{\text{HMM}}(\bar{X}^0;\cdot)] - \mathbb{E}[\bar{F}_{\text{HMM}}(\bar{X}^0)]\right). \tag{19}$$

From equation (19), we observe that the variance-reduced HMM estimator $\bar{F}_{\text{HMM}}(\bar{X}^N)$ does not introduce an additional bias compared to the original estimator $\hat{F}_{\text{HMM}}(\bar{X}^N;\omega_N)$ as soon as the initial variance-reduced estimator $\bar{F}_{\text{HMM}}(\bar{X}^0)$ is unbiased with respect to $\hat{F}_{\text{HMM}}(\bar{X}^0;\omega_0)$. Furthermore, in the limit of $M,\tau\to\infty$ and δt constant we obtain:

$$\lim_{\substack{M\to\infty\\ M\delta t=\tau}} \mathbb{E}[\bar{F}_{\text{HMM}}(\bar{X}^N)] = \mathbb{E}_{\hat{\mu}^\infty_{\bar{X}^N}}[f(\bar{X}^N,\cdot)] - \left(\mathbb{E}_{\hat{\mu}^\infty_{\bar{X}^0}}[f(\bar{X}^0,\cdot)] - \mathbb{E}_{\bar{\mu}^\infty_{\bar{X}^0}}[f(\bar{X}^0,\cdot)]\right), \tag{20}$$

in which $\hat{\mu}^\infty_{\bar{X}^N}$ denotes the invariant measure of the Markov chain with fixed value $\bar{X}^N$ of the slow variable, and $\bar{\mu}^\infty_{\bar{X}^0}$ corresponds to the invariant measure of samples used in the initial estimation.

Second, we study the statistical error which is described by the variance of the estimator:

$$\text{Var}[\bar{F}_{\text{HMM}}(\bar{X}^N)] = \text{Var}\left[\hat{F}_{\text{HMM}}(\bar{X}^N;\omega_N) - \left(\hat{F}_{\text{HMM}}(\bar{X}^{N-1};\omega_N) - \bar{F}_{\text{HMM}}(\bar{X}^{N-1})\right)\right]. \tag{21}$$

By using Taylor series expansions, it can be shown that equation (21) is further calculated as:

$$\text{Var}\left[\bar{F}_{\text{HMM}}(\bar{X}^N)\right] \approx \text{Var}\left[\left(1+\frac{\Delta t}{M}\sum_{m=1}^{M}\left(f_x^m + f_y^m\sum_{i=1}^{m-1}\left(B_N^i\prod_{j=i+1}^{m-1}A_N^j\right)\right)\right)\bar{F}_{\text{HMM}}(\bar{X}^{N-1})\right], \tag{22}$$

in which we used the shorthand notation f_x^m and f_y^m to denote the partial derivative of the function f with respect to x and y, respectively, evaluated at $(\bar{X}^N, y_N^{N,m})$ and we introduced:

$$A_N^m = 1+\frac{\delta t}{\varepsilon}g_y^m + \sqrt{\frac{\delta t}{\varepsilon}}\beta_y^m\cdot\xi_N^m, \qquad B_N^m = \frac{\delta t}{\varepsilon}g_x^m + \sqrt{\frac{\delta t}{\varepsilon}}\beta_x^m\cdot\xi_N^m.$$

4.2 Pathwise Convergence

The variance-reduced HMM solution of the macroscopic equation is written as:

$$\begin{aligned}\bar{X}^N &= \bar{X}^{N-1} + \Delta t\bar{F}_{\text{HMM}}(\bar{X}^{N-1})\\ &= \bar{X}^0 + \Delta t\sum_{n=0}^{N-1}\bar{F}_{\text{HMM}}(\bar{X}^n).\end{aligned} \tag{23}$$

Using expression (18) followed by taking expectations of both sides of equation (23) we obtain:

$$\begin{aligned}\mathbb{E}[\bar{X}^N] = \bar{X}^0 &+ \Delta t\left(\mathbb{E}[\bar{F}_{\text{HMM}}(\bar{X}^0)] + \sum_{n=1}^{N-1}\mathbb{E}[\hat{F}_{\text{HMM}}(\bar{X}^n;\omega_n)]\right)\\ &+ \Delta t(N-1)\left(\mathbb{E}[\bar{F}_{\text{HMM}}(\bar{X}^0)] - \mathbb{E}[\hat{F}_{\text{HMM}}(\bar{X}^0;\omega_1)]\right).\end{aligned} \tag{24}$$

Next, we examine how paths behave by looking at the error E^N at time instance t^N between the variance-reduced HMM solution $\bar{X}^N$ and the exact solution $X(t^N)$. To show convergence for a fixed time instance $t^N = N\Delta t$, we calculate an upper bound for the mean-squared error which tends to 0 for $\Delta t \to 0$. To that end, we introduce the forward Euler solution of the averaged equation (2) at time instance t^N using the exact expression of the function F, which we denote by X^N. This leads to:

$$\begin{aligned} E^N &= |\bar{X}^N - X(t^N)| \\ &\leq |\bar{X}^N - X^N| + |X^N - X(t^N)| \\ &\leq |\bar{e}^N_{\text{HMM}}| + O(\Delta t). \end{aligned} \tag{25}$$

The first term corresponds to the variance-reduced HMM error, while the second term is a classical result of a forward Euler discretization. The former can be further calculated as:

$$\bar{e}^N_{\text{HMM}} = \Delta t \sum_{n=0}^{N-1} e_s(\bar{F}_{\text{HMM}}(\bar{X}^n)) + \Delta t \sum_{n=0}^{N-1} e_b(\bar{F}_{\text{HMM}}(\bar{X}^n)) + \Delta t \sum_{n=0}^{N-1} \left(F(\bar{X}^n) - F(X^n)\right). \tag{26}$$

in which we introduced the following shorthand notation for the systematic and statistical error:

$$e_b(\bar{F}_{\text{HMM}}(\bar{X}^n)) = \tilde{\bar{F}}_{\text{HMM}}(\bar{X}^n) - F(\bar{X}^n), \qquad e_s(\bar{F}_{\text{HMM}}(\bar{X}^n)) = \bar{F}_{\text{HMM}}(\bar{X}^n) - \tilde{\bar{F}}_{\text{HMM}}(\bar{X}^n). \tag{27}$$

To bound the error, we look at the expected value of the error squared. Using Lipschitz-continuity of F with Lipschitz constant L and applying Gronwall's inequality, the following error bound is obtained [3]:

$$\mathbb{E}\left[|\bar{e}^N_{\text{HMM}}|^2\right] \leq 3T\Delta t S^N \exp\left(3T^2L^2\right), \tag{28}$$

in which we used $N\Delta t = T$ and

$$S^N = \sum_{n=0}^{N-1} \mathbb{E}\left[\left|e_b(\bar{F}_{\text{HMM}}(\bar{X}^n))\right|^2 + \left|e_s(\bar{F}_{\text{HMM}}(\bar{X}^n))\right|^2\right]. \tag{29}$$

This last bound at a fixed time instance $t^N = N\Delta t$ indeed converges to 0 in the limit $\Delta t \to 0$ which concludes the convergence of the method.

5 Numerical Results

We now apply the proposed variance reduction method to a linear and nonlinear model problem in sections 5.1 and 5.2, respectively.

5.1 Linear System

Let us first demonstrate the variance reduction technique when applied to the linear model problem given in equation (4). We begin by applying the HMM technique without variance reduction, in which we approximate the reduced evolution of the slow variable in equation (6) which is in turn an approximation of the slow variable's true evolution described in system (4). We compute the solution for $t \in [0, 1]$ using initial conditions $\hat{X}^0 = x_0 = 1$ and $y_0 = 1$. The system parameters in equation (4) are as follows: $\lambda = -10$, $p = 4$, $q = 0.5$ and $A = 1$.

In every iteration of the HMM method, we generate $M = 50$ samples by simulating the fast equation of system (4), while keeping the current value of the slow variable fixed. For stability, the simulation time step δt in the Euler-Maruyama discretization is chosen as $\delta t = \varepsilon$ with $\varepsilon = 10^{-3}$. The forward Euler time step used in the discretization of the macroscopic equation (2) is fixed as $\Delta t = 0.02$. The evolution of the variables X and F and their variance are shown by blue circles in figure 2. Clearly, the statistical error dominates, thus justifying the need for variance reduction.

Before applying the variance reduction technique based on control variables introduced in section 3.2, we first specify how the function F is estimated in the initial forward Euler step of the macroscopic equation shown in equation (14). If this first estimation of F is performed using its exact expression as calculated in equation (6), the variance-reduced HMM estimator is completely variance-free, thus leading to a deterministic estimator. This result is briefly illustrated: in the second forward Euler step of the macroscopic equation, combining equation (15) with (13), the variance of the variance-reduced estimator is obtained as follows:

$$\begin{aligned}\mathrm{Var}[\bar{F}_{\mathrm{HMM}}(\bar{X}^1)] &= \mathrm{Var}\left[\hat{F}_{\mathrm{HMM}}(\bar{X}^1;\omega_1) - \left(\hat{F}_{\mathrm{HMM}}(\bar{X}^0;\omega_1) - F(\bar{X}^0)\right)\right] \\ &= \frac{p^2}{M^2}\mathrm{Var}\left[\sum_{m=1}^{M}(y_m^1 - y_m^0)\right]. \end{aligned} \tag{30}$$

Given that $F(\bar{X}^0)$ is evaluated using its exact form, in equation (30) we used that $\bar{X}^0$, $\bar{X}^1$ and $F(\bar{X}^0)$ are all deterministic quantities. To calculate the sum in equation (30), we first subtract the equations of the two Markov chains, as given in equation (11), from each other resulting in:

$$y_1^{1,m+1} - y_1^{0,m+1} = \left(1 - A\frac{\delta t}{\varepsilon}\right)(y_1^{1,m} - y_1^{0,m}) + \frac{\delta t}{\varepsilon}q(\bar{X}^1 - \bar{X}^0) + \sqrt{\frac{\delta t}{\varepsilon}}(\xi_1^m - \xi_1^m). \tag{31}$$

Since we are using the same Brownian path for both Markov chains, the stochastic part of equation (31) cancels out exactly. Therefore, this difference between Markov chain generated samples is completely deterministic and its variance is zero. This continues to hold for all following forward Euler steps. The results of the variance-reduced estimator are shown in figure 2. Green circles represent a simulation with exact initialization, while red and cyan circles are used for simulations with an HMM estimated initial step using sample sizes of 10^3 and 10^4, respectively, for which only the variance of X and F are shown for clarity.

5.2 Nonlinear System

As a second model problem, we consider the nonlinear stochastic multiscale system given in equation (7). We begin by applying the HMM procedure without variance reduction. We calculate the solution for $t \in [0, 2]$ using initial conditions $\hat{X}^0 = x_0 = y_0 = 0.5$. In the HMM method, we use $M = 50$ samples by iterating over the Euler-Maruyama scheme for the fast dynamics with time step $\delta t = \varepsilon$ and $\varepsilon = 10^{-3}$. We fix the forward Euler time step as $\Delta t = 0.05$. The results are depicted by blue circles in figure 3. The HMM method now fails to capture the correct solution path: a clear bias has developed for both X and F which reappears over repeated experiments. This is attributed to the difference between the true measure $\mu_x^\infty(y)$ of the fast dynamics and the approximate measure $\hat{\mu}_x^\infty(y)$ of the Markov chain which possesses the following invariant mean $\hat{m}_\infty$ and variance $\hat{\sigma}_\infty^2$:

$$\hat{m}_\infty = x, \qquad \hat{\sigma}_\infty^2 = \frac{1}{2 - \dfrac{\delta t}{\varepsilon}}. \tag{32}$$

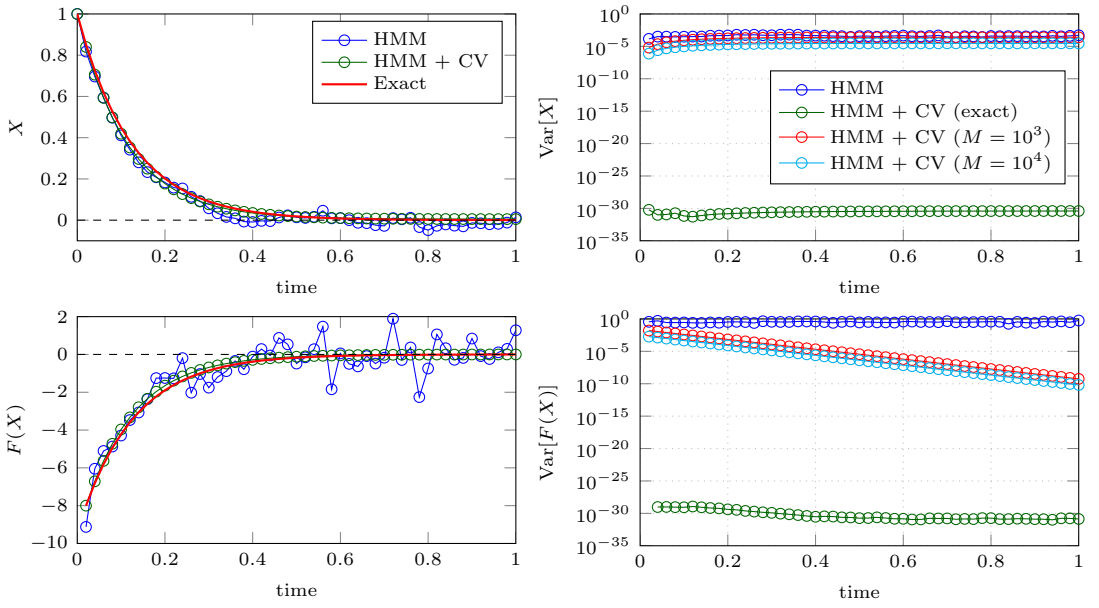

Figure 2: Left: X and F of the HMM method applied to the linear model problem (4) with and without variance reduction (green and blue circles, respectively). The red line represents the exact solution. Right: variance of X and F for HMM. Blue circles: no variance reduction; green circles: variance reduction and exact first step; red and cyan circles: variance reduction and estimated first step using $M = 10^3$ and $M = 10^4$ samples, respectively (color online).

Comparing this result with (8), we conclude that the measures $\mu_x^\infty(y)$ and $\hat{\mu}_x^\infty(y)$ have the same invariant mean but different invariant variance due to the Euler-Maruyama discretization of the fast equation. Moreover, this difference disappears for $\delta t \to 0$. However, when choosing $\delta t = \varepsilon$ as above, the HMM method yields an estimator for the macroscopic equation with right hand side $F(X) = -(X + X^2 + 1)$ instead of the desired equation given in (9).

To correct this difference in invariant measure, we supplement the HMM method with the Metropolis-Hastings algorithm [6]. This algorithm introduces an accept-reject procedure for every sample generated by the Markov chain such that the obtained ensemble follows the desired invariant distribution. The results of the HMM method with the Metropolis-Hastings algorithm are shown by cyan circles in figure 3. We observe that the bias has vanished and the statistical error of the estimator has become the dominant error source.

Unfortunately, when combining the Metropolis-Hastings algorithm with the variance-reduced HMM estimator, we lose strong correlation between the HMM estimators $\hat{F}_{\text{HMM}}(\bar{X}^n; \omega_n)$ and $\hat{F}_{\text{HMM}}(\bar{X}^{n-1}; \omega_n)$ in equation (13) since samples can be rejected in different places in both ensembles. To resolve this, we instead use two classical HMM estimators without the Metropolis-Hastings extension in equation (13), each producing a bias. However, since both contain the same bias, subtraction yields a result of $O(\Delta t)$. In the simulations, we used the exact expression of F in the initial forward Euler step of the averaged equation. The results are shown by green circles in figure 3. We remark that the variance buildup of the variance-reduced estimator on the right hand side plots can be countered by occasionally reinitializing the estimator.

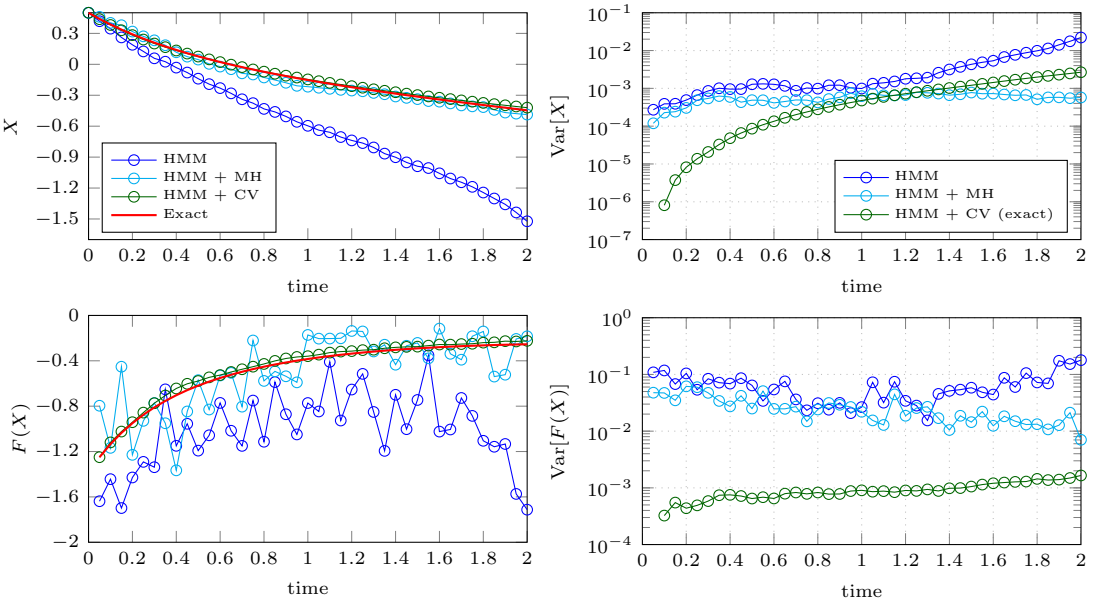

Figure 3: Left: X and F of the HMM method applied to the nonlinear model problem (7) with and without variance reduction (green and blue circles, respectively). The red line represents the exact solution. Right: variance of X and F for HMM. Blue circles: no variance reduction; green circles: variance reduction and exact first step; red and cyan circles: variance reduction and estimated first step using $M = 10^3$ and $M = 10^4$ samples, respectively (color online).

6 Conclusions

We proposed a novel variance reduction strategy based on control variables for simulating the averaged equation of a stochastic slow-fast system. We analyzed convergence of the proposed estimator and applied it to a linear and nonlinear model problem. The numerical tests showed a significant reduction of variance. For future work, it is interesting to apply the variance reduction technique to metastable systems [1] and systems containing oscillatory functions.

References

[1] Maria Bruna, S. Jonathan Chapman, and Matthew J. Smith. Model reduction for slow-fast stochastic systems with metastable behaviour. *Journal of Chemical Physics*, 140(17):1–23, 2014.

[2] Radek Erban, S. Jonathan Chapman, Ioannis G. Kevrekidis, and Tomáš Vejchodský. Analysis of a stochastic chemical system close to a sniper bifurcation of its mean field model. *SIAM Journal on Applied Mathematics*, 70(3):984–1016, 2009.

[3] Dror Givon, Ioannis G. Kevrekidis, and Raz Kupferman. Strong convergence of projective integration schemes for singularly perturbed stochastic differential systems. *Communications in Mathematical Sciences*, 4(4):707–729, 2006.

[4] Dror Givon, Raz Kupferman, and Andrew Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):55–127, 2004.

[5] Desmond J. Higham. An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Review*, 43(3):525–546, jan 2001.

[6] Tony Lelièvre, Gabriel Stoltz, and Mathias Rousset. *Free energy computations: A mathematical perspective*. World Scientific, 2010.

[7] Tiejun Li, Assyr Abdulle, and Weinan E. Effectiveness of implicit methods for stiff stochastic differential equations. *Communications in Computational Physics*, 3(2):295–307, 2008.

[8] Ward Melis and Giovanni Samaey. Variance-reduced HMM for stochastic slow-fast systems. In preparation, 2015.

[9] Anastasia Papavasiliou and Ioannis G. Kevrekidis. Variance reduction for the equation-free simulation of multiscale stochastic systems. *Multiscale Modelling and Simulation*, 6(1):70–89, 2007.

[10] Grigorios A. Pavliotis and Andrew Stuart. *Multiscale methods: averaging and homogenization*. Springer Science & Business Media, 2008.

[11] Mathias Rousset and Giovanni Samaey. Individual-Based Models for Bacterial Chemotaxis in the Diffusion Asymptotics. *Mathematical Models and Methods in Applied Sciences*, 23(11):2005–2037, 2013.

[12] Eric Vanden-Eijnden. Numerical techniques for multi-scale dynamical systems with stochastic effects. *Communications in Mathematical Sciences*, 1(2):385–391, 2003.